Supplementary information

Table of Contents

Table of Figures

1 Algorithm description

1.1 Data pre-processing

SPOT has been preinstalled with **three datasets**, one single cell RNAseq dataset describing the whole *Plasmodium* life cycle, a bulk RNAseq dataset containing information about human organs in several developmental stages and a bulk RNAseq dataset from a SARS-CoV-2 infected cell line (Cardoso-Moreira, et al., 2019; Howick, et al., 2019; Wyler, et al., 2021).

For spot analysis of the ***Plasmodium* life cycle**, TMMlog normalized single cells were assigned by ShortenedLifestage4 (Howick, et al., 2019) and averaged to obtain a single expression value for every developmental stage.

Human organ TPM values were also condensed by averaging between multiple replicates. Since the similarity of prenatal developmental stages is high, only 3 prominent stages (4, 10, 20 weeks post conception) were kept for further analysis.

TPM values of the **SARS-CoV-2 infected cell line** were calculated by the gene length specified by the authors (Wyler, et al., 2021). The pre-loaded samples show one duplicate of every timepoint per series.

For *spot* and correlation ranking, data is **scaled by standardization** to obtain comparability between datasets. To lower the influence of outliers every standardized value higher than 4 was set to 4 and every standardized value below -1 was set to -1.

In contrast to *spot*, for **differential expression analysis** raw counts were used and filtered for cells having expression in more than 3 features.

1.2 SPOT

The *spot* algorithm consists of two factors: (i) the difference between the weighted mean of selected (slider values > 0) or unselected entities (slider values = 0) and (ii) the difference between 1 and the mean of unselected entities. For an optimal result, both factors should be as high as possible.

The first factor measures the distance between the selected and unselected entities, by calculating the difference of their mean expression values. To give the user the possibility to adapt the results to his or her needs, the mean value of the selected entities can be weighted. For this purpose, individual entities can be assigned to the mean value several time by adjusting the slider values, thus increasing the influence of the respective entity. The higher the slider value, the

more frequently the variable is assigned to the mean, the lower the slider value, the less frequently the variable is assigned. A showcase how this weighting can improve the results is shown in Figure S1.

The second factor measures how close the expression values of the unselected entities are to zero (desired in this approach). The closer they are, the less gets subtracted from one and therefore the second factor increases. Due to the scaling it is also possible to obtain negative values, for the mean.

Taken together both factors result in the *spot* score which is therefore a measure for the proximity to the user defined input. According to the definition and data pre-processing *spot*-scores can lie between -1 and 10, however SPOT displays only positive ones. The highest values obtained in the pre-loaded datasets are around 8.

With N as the number of entities or columns, α_i as the weights from the slider values and c_i as the respective expression values, *spot* is defined as:

$$I := \{1, \dots, N\}$$

$$I_s := \{1 \leq i \leq N \mid column\ i\ selected\}$$

$$n_s = |I_s|$$

$$\overline{sc} = \frac{\sum_{i \in I_s} \alpha_i c_i}{\sum_{i \in I_s} \alpha_i}$$

$$\overline{uc} = \frac{1}{N - n_s} \sum_{i \in I \setminus I_s} c_i$$

$$spot = (\overline{sc} - \overline{uc})(1 - \overline{uc})$$

1.3 Correlation

To enable complex profiles and medium values between high and low expression, we implemented an option in the user interface simply calculating the Pearson correlation between the user defined profile and the genes in the dataset. As a result of the cutoff (see chapter 1.1) the highest values in the datasets are equal to 4, while the lowest generally lie close to 0. The POI defined by the slider values is matched with these normalized values, which allows searching for complex profiles with several subcategories between very low and high expression. The results of the ranking after the correlation calculation when only one slider is moved can be seen in Figure S2. Compared to SPOT, the mean expression values in unselected columns are higher than in Spot or DEA approaches (Figure S4).

Figure S1: Profile tuning highlights stage specifically expressed genes

(a) A SPOT search for genes highly expressed in ookinetes, oocysts, sporozoites and liver stage parasites; with all sliders of the respective stages set to 1 (all stages count equally for the weighted sum). Consistent with this POI, the top genes, shown as a dot plot on the right, show high expression in ookinetes and oocysts and slightly lower expression in sporozoites and liver stages. Blood stages show almost no expression as entered in the POI. **(b)** Searching for the same stages as in **(a)**, but with different weights for each stage. Expression in liver stage parasites counts double in the weighted sum, while expression in sporozoites counts once and in ookinetes and oocysts only half. As a result, the top genes have high expression in liver stages, medium expression in sporozoites, ookinetes and oocysts and again almost no expression in blood stages. **(c)** A search similar to **(b)** but setting the slider for expression in sporozoites to two and the slider for liver stages to one. This POI yields top genes with strong expression in sporozoites, while they have medium high values in liver stages, ookinetes and oocysts. Again, blood stages show almost no expression. The circle size in the dot plot corresponds to the number of single cells in which the respective gene is detected; the color of the dot represents the average expression as indicated by the heat bar on the right.

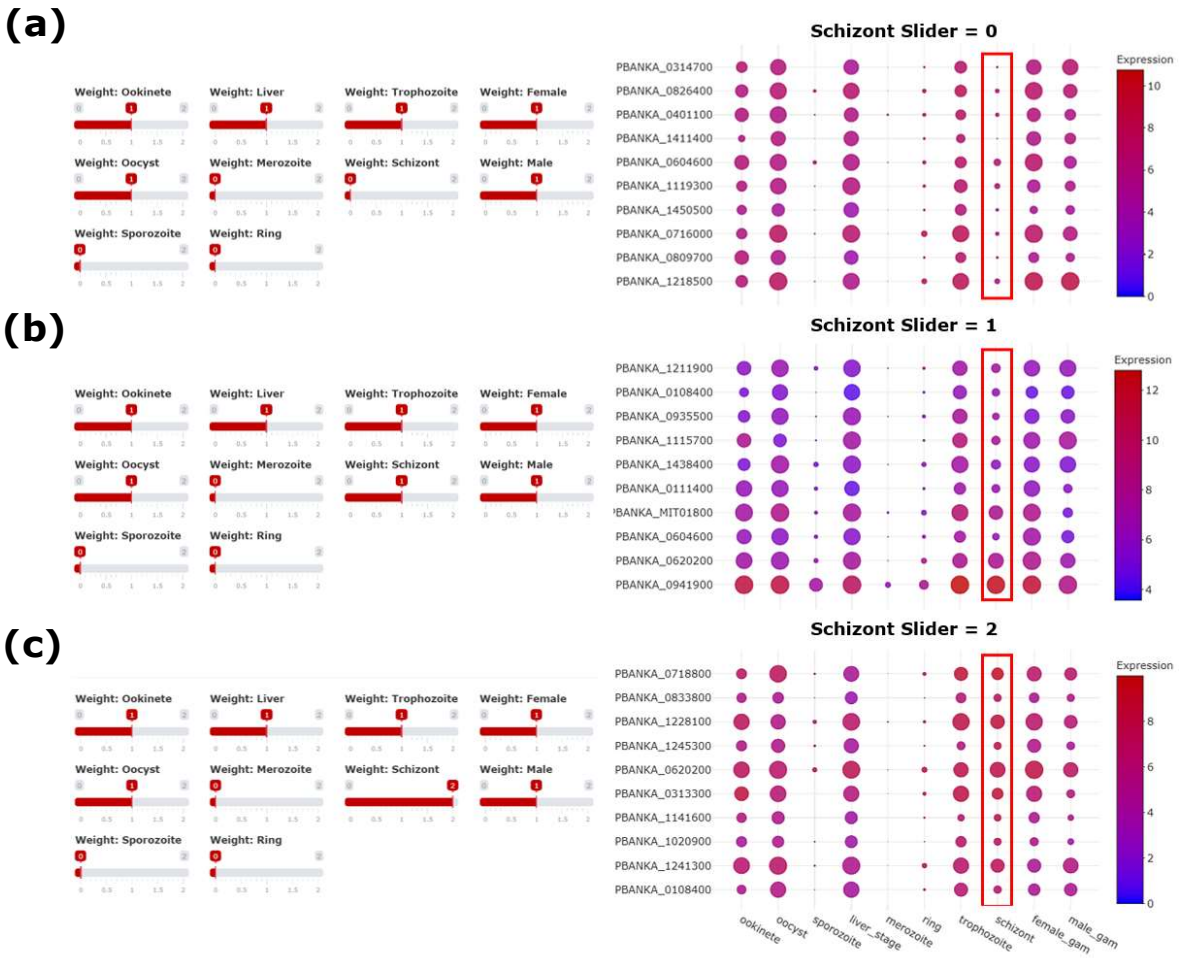

Figure S2: Highlighting medium expression for specific stages

(a) A correlation search for high expression in 6 of the 10 displayed lifecycle stages of *P. berghei* (Ookinete, Oocyst, Liver stage, Trophozoite, Schizont, Female and Male gametocytes). The Schizont slider is set to 0, which indicates that in gene expression profiles that should score high in the ranking, no expression in the schizont stage is desired. **(b)** A similar search to **(a)** with the schizont slider set to 1. As can be seen in the red rectangle, the expression in this stage increased substantially, while the others stay at a level comparable to **(a)**. **(c)** A similar search as in **(a)** with the schizont slider set to 2. The gene expression in the displayed top ranked genes is higher in the schizont stage as well as in the other selected stages, compared to **(a)** and **(b)**. The circle size in the dot plot corresponds to the number of single cells in which the respective gene is detected; the heat bar on the right displays the average expression.

1.4 DEA

Differential expression analysis is performed with the help of the R packages Seurat, MAST and DESeq2 (Hao *et al*., 2020; McDavid *et al.* 2020; Love *et al.,* 2014). Single cell datasets such as the data derived from the Malaria Cell Atlas are loaded to a Seurat object and identities are specified according to the ShortenedLifestage4 clustering (Howick, et al., 2019). The Seurat function FindMarkers then compares two groups of entities entered by the user with different test methods (see Section 2). In the output table the Bonferroni corrected p-value and the log fold change are displayed. If desired by users, implementation of further analysis methods is readily feasible.

2 Comparisons of algorithms and tools

2.1 Algorithms

Since there are several approaches to detect significantly up-regulated genes, we compared the *spot* and correlation algorithm with the state-of-the-art methods of differential analysis in speed and accuracy.

The speed was tested on 3 example predictions, which are described in the further sections. We measured the time between input of the profile and output of the results with several CPUs. The results shown in Figure S3 depict an inverse association between the simplicity of the ranking method and the speed. While the *spot* and correlation ranking finished the analysis within 3 seconds, the Wilcoxon Rank Sum test outperformed the other DEA methods such as DESeq2 and Mast.

Please note that performance may vary between different operating system/browser combinations.

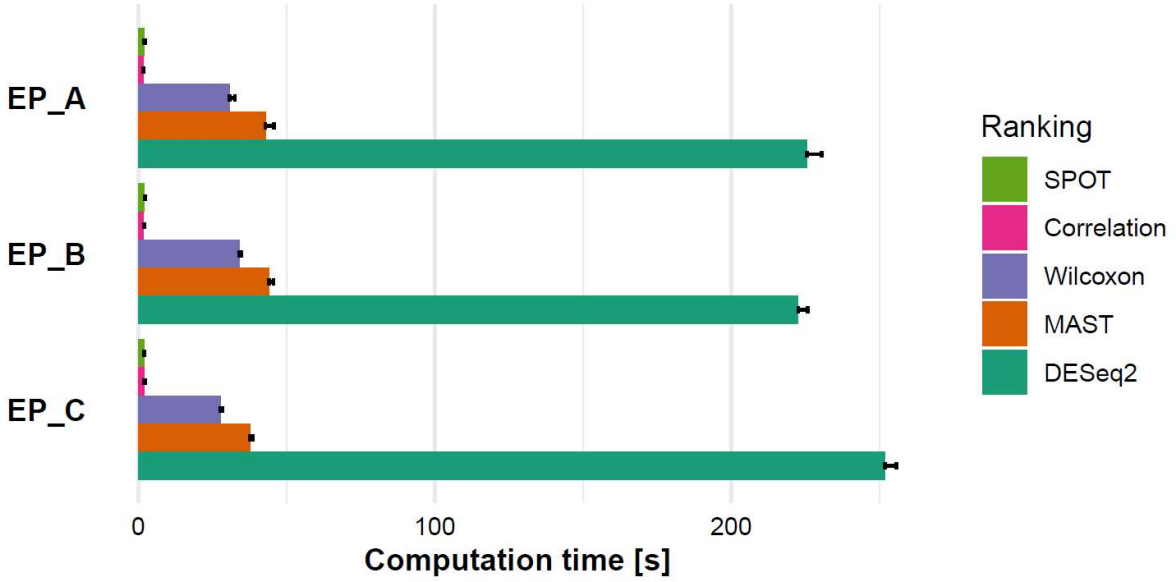

Figure S3: *spot* algorithm outperforms DEA approaches in speed

spot, Correlation and DEA methods were benchmarked for time between input in the control panel and output as table in the web version of the tool. While SPOT and correlation analysis stayed within 3 seconds per calculation, DEA methods ranged from approx. 30 to 250 seconds.

Accuracy tests were performed by measuring the mean expression of the best ranked candidates in selected and unselected entities of the example predictions (Figure S4). In general, the more entities are selected, the closer the values between the selected and unselected values tend to be in the example predictions. The results in the three predictions show similar accuracy of all methods with only subtle differences.

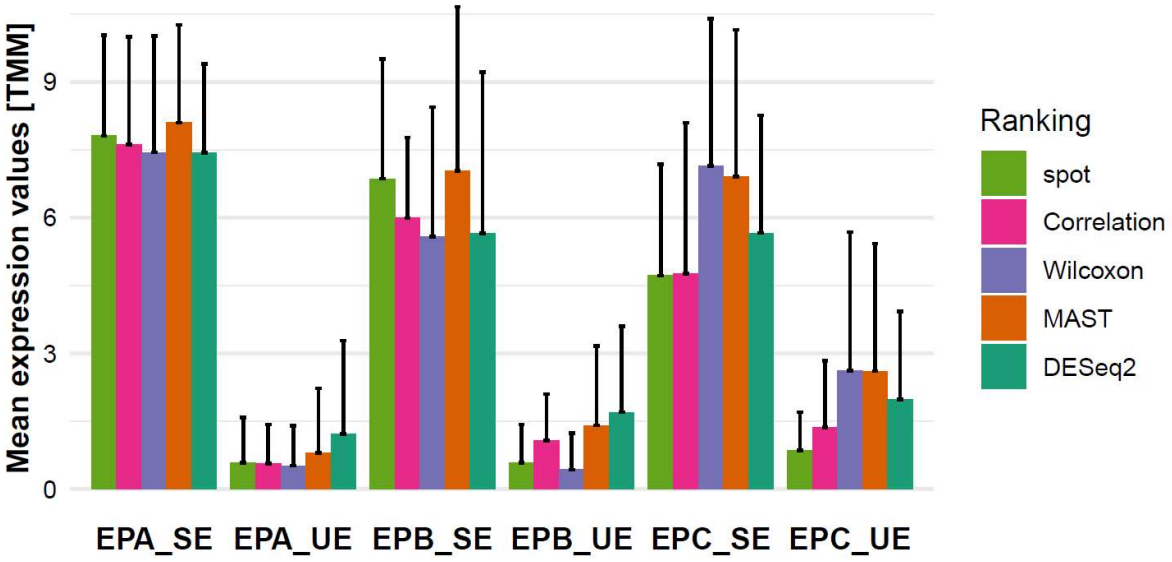

Figure S4: Comparison of selected and unselected entities reveal best results for *spot*

Ranking accuracy in example predictions (EP) was determined through mean expression calculation of selected (SE) and unselected (UE) entities. There is overall similar accuracy revealed with MAST seemingly the best algorithm to get genes with high expression in selected entities; *spot* and the Wilcoxon test score best in unselected entities.

Since the Wilcoxon Rank Sum test is the fastest DEA method (Figure S3), has best accuracy amongst unselected entities (Figure S4) and has also shown satisfactory results in the literature (Soneson and Robinson, 2018) it is used as the default in the web tool.

2.2 Tools

Since there are multiple tools available for DEA (Ge, et al., 2018; Reyes, et al., 2019), we would like to point out the strengths and weaknesses of the approaches (Table S1). While both approaches show satisfactory results in ranking of genes (Figure S4), there are differences in speed, visualization and interface design. Although the multitude of available visualization methods gives iDEP and GENAVi an advantage over SPOT, easier handling and faster calculations make SPOT a viable alternative. SPOT might well be the better choice for users more interested in a brief overview, than in an in-depth analysis. Thus, both approaches have their advantages and disadvantages and give users the opportunity to choose according to their needs.

Table S1: Comparison of web-tools for expression analysis

Parameter	GENAVi	iDEP	SPOT
Accuracy	Good	Good	Okay
Speed	Moderate	Moderate	Good
Visualization	Good	Okay	Okay
User interface	Okay	Okay	Good

3 Example prediction A: *Plasmodium* liver stages

3.1 Purpose

There is still no efficient malaria vaccine available (2015; Duffy and Patrick Gorres, 2020). Different approaches are currently being explored aiming to target parasites in the disease-causing blood stages or during transmission to and from the mosquito. Transmission blocking vaccines focus e.g. on generating antibodies against proteins present on the surface of gametes, which can block the transmission of *Plasmodium* to the mosquito (Carter and Chen, 1976; Gwadz, 1976), while blood stage vaccines aim e.g. at blocking parasite entry into red blood cells. In addition to such subunit vaccines also attenuated parasites are explored. Attenuated parasites are most often generated by genetic modification of genes functional during liver stage development, which can lead to a developmental arrest in hepatocytes (Kumar, et al., 2016; Mueller, et al., 2005). The specific genes are usually determined by differential expression analysis (Kaiser, et al., 2004; Matuschewski, et al., 2002). However, none of the approaches have yet succeeded in inducing sufficient protective immune responses (Duffy and Patrick Gorres, 2020). Here we present the results of a SPOT search for genes highly expressed exclusively in the liver stages. Since the genes with such expression profiles are well known (Caldelari, et al., 2019; Stanway, et al., 2019), this prediction can serve as proof-of-concept.

3.2 Results

The SPOT search for genes with high expression during liver stage development revealed a top ten of genes shown in **Table S2**. Liver specific proteins 1 and 2 appear in the first 3 ranks/positions, while 3 genes with unknown functions appear in the top 6. For all known genes except for the gametocyte specific protein and pyruvate dehydrogenase E1 component subunit beta, literature suggests strong liver stage expression (Ishino, et al., 2009; Orito, et al., 2013; Stanway, et al., 2019; Vaughan, et al., 2009). For PBANKA1003900, previously annotated as gametocyte-specific protein, recent analysis led to the suggestion of renaming it liver-specific protein due to its liver specific expression profile (Caldelari, et al., 2019; Deligianni, et al., 2018).

Table S2: Results from *spot* ranking of liver specific genes

P. berghei GeneID	Gene product description	*spot* score
PBANKA_1024600	liver specific protein 1	8,01
PBANKA_0519500	conserved Plasmodium protein, unknown function	7,75
PBANKA_1003000	liver specific protein 2	7,19
PBANKA_1003900	gametocyte-specific protein	6,63
PBANKA_0518900	conserved Plasmodium membrane protein, unknown function	5,80
PBANKA_1462600	conserved Plasmodium protein, unknown function	3,83
PBANKA_0505000	dihydrolipoamide acyltransferase, putative	3,62
PBANKA_1310100	pyruvate dehydrogenase E1 component subunit beta, putative	3,58
PBANKA_1125100	3-oxoacyl-acyl-carrier protein synthase, putative	3,08
PBANKA_1303600	leucine carboxyl methyltransferase, putative	3,08

3.3 Comparison to DEA

To check the quality of the *spot* ranking, we compared the best performing genes of the *spot* ranking with those of the DEA methods. While the first 100 genes in the *spot* ranking are significantly upregulated according to the Wilcoxon test, 14 of the top 20 overlap in both rankings (Figure S5).

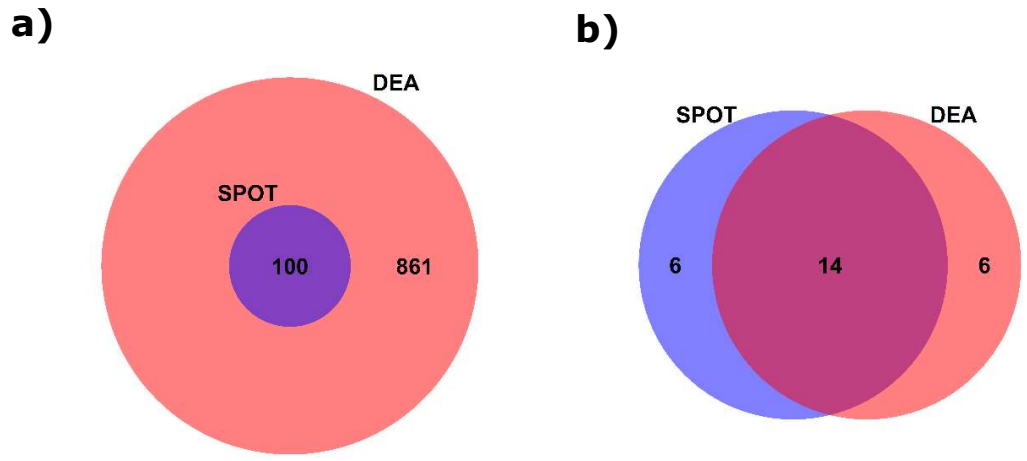

Figure S5: First hundred genes ranked in SPOT liver stage ranking are differentially expressed

(a) 961 *P. berghei* genes are upregulated exclusively in the liver (p-value < 0,001). The first 100 genes of the *spot* ranking overlap with these results. **(b)** Two thirds (14) of the top 20 genes (for DEA ranked by p-value) of both tests overlap.

The best performing candidates of *spot* and Wilcoxon ranking are displayed in Figure S6. They reveal strong expression in the *Plasmodium* liver stage as well as weaker expression in trophozoites and oocysts. While the genes performing best in the *spot* ranking show higher expression values in oocysts, the Wilcoxon test ranked genes have more expression in trophozoites. As shown in Figure S6, there is less overall expression in unselected columns of the Wilcoxon ranking, while the *spot* ranked genes have higher expression in the liver stage.

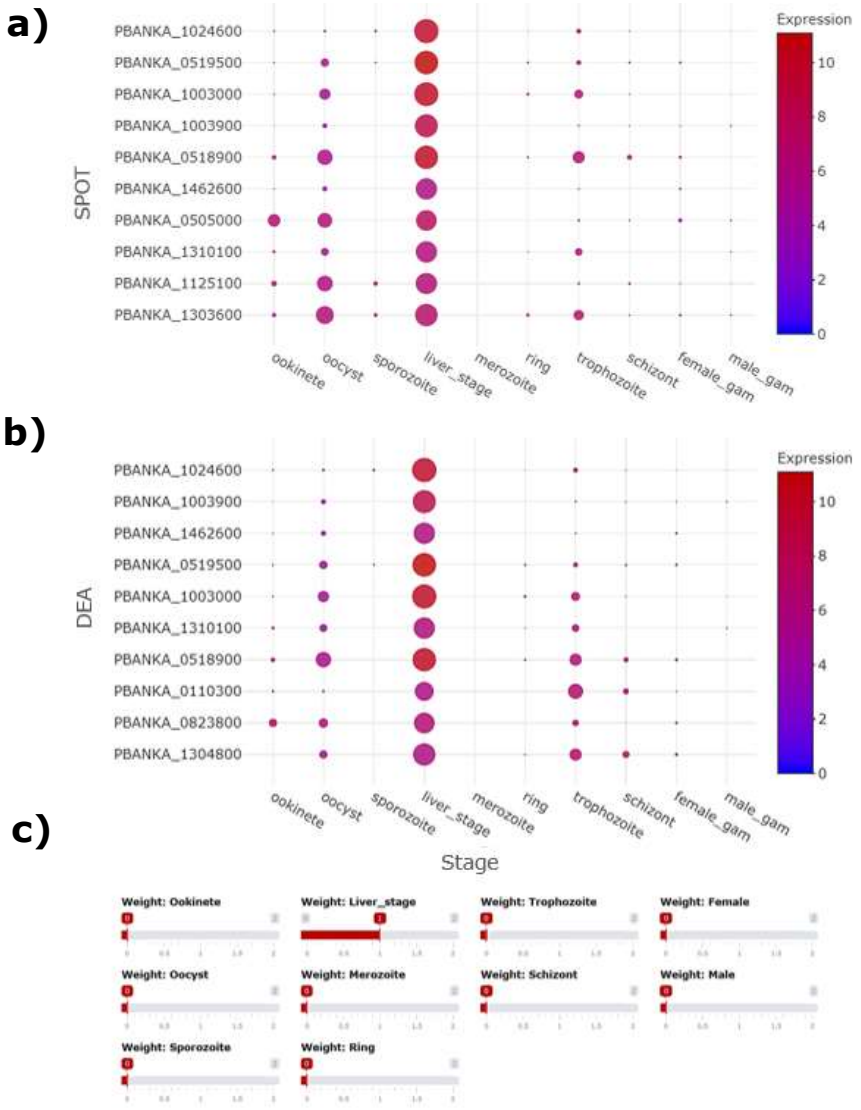

Figure S6: Comparison of SPOT and DEA reveal higher off-target effects in *spot* ranking

The 10 best ranked genes from *spot* **(a)** and Wilcoxon ranking **(b).** While the genes of the Wilcoxon ranking have slightly lower values in the unselected columns, *spot* ranked genes have higher values in the liver stage. **(c)** The slider constellation of SPOT to obtain the results shown in **(a)** and **(b)**.

4 Example prediction B: *Plasmodium* sexual stages

4.1 Purpose

Drug development against malaria focuses mainly on the blood stages of malaria, as these cause the symptoms of the disease and are easily accessible in the laboratory. Since the eradication of malaria has come into the focus of drug design, drugs are also developed against sexual stages to prevent transmission to the mosquito or against liver stages to prevent development of disease-causing blood stages (Kappe, et al., 2010). Drugs targeting the liver stages are routinely used to cure *Plasmodium vivax* infections (Flannery, et al., 2018). However, there are no specific drugs available that target gametocytes. Identifying additional target proteins is therefore of interest (The malERA Consultative Group on Drugs, 2011). Here were present a SPOT search for genes only expressed in sexual blood stages, leaving a high probability for functionality in these stages.

4.2 Results

Table S3: Results from *spot* ranking for sexual stage specific genes

P. Berghei GeneID	Gene product description	*spot* score
PBANKA_1329100	plasmepsin VIII	4,95
PBANKA_0600600	NIMA related kinase 3, putative	4,95
PBANKA_1432200	male development gene 1	4,79
PBANKA_1429100	conserved Plasmodium protein, unknown function	4,78
PBANKA_1449000	microgamete surface protein MiGS, putative	4,49
PBANKA_1361600	E1-E2 ATPase, putative	4,25
PBANKA_1038200	nuclear formin-like protein MISFIT	4,22
PBANKA_0812600	conserved Plasmodium protein, unknown function	4,18
PBANKA_1359600	6-cysteine protein	4,12
PBANKA_1109600	conserved Plasmodium protein, unknown function	4,07

Sexual stage specific *spot* ranking revealed 3 unknown proteins as well as a wide variety of enzymes, surface proteins and proteins involved in osmiophilic body formation. *P. falciparum* plasmepsin VIII is known to be active specifically in gametocytes while the p48/45 protein (PBANKA_1359600) is a well characterized surface protein and still in consideration as a target for transmission blocking vaccines (Acquah, et al., 2019; Jiang, et al., 2020; Lee, et al., 2020; van Dijk, et al., 2001; Weißbach, et al., 2017).

Two other interesting candidates are associated with the emergence of osmiophilic bodies, specialized vesicles essential for parasite egress from blood cells and another one was shown to be important in the cell cycle in ookinetes and oocysts (Bushell, et al., 2009; Kehrer, et al., 2016; Ponzi, et al., 2009).

It is therefore hard to predict the function or localization of the 3 unknown genes. However, the interesting phenotypes of proteins with a similar transcription profile suggests that these proteins may have a functional role in mosquito infection as well.

4.3 Comparison to DEA

Similar to the previous section, all genes ranked in the Top 100 of *spot* algorithm are significantly upregulated exclusively in the selected sexual stages (Figure S7 a). In contrast to the expression profiling of liver stages, however, approximately only half of the genes in the top 20 still overlap (Figure S7 b).

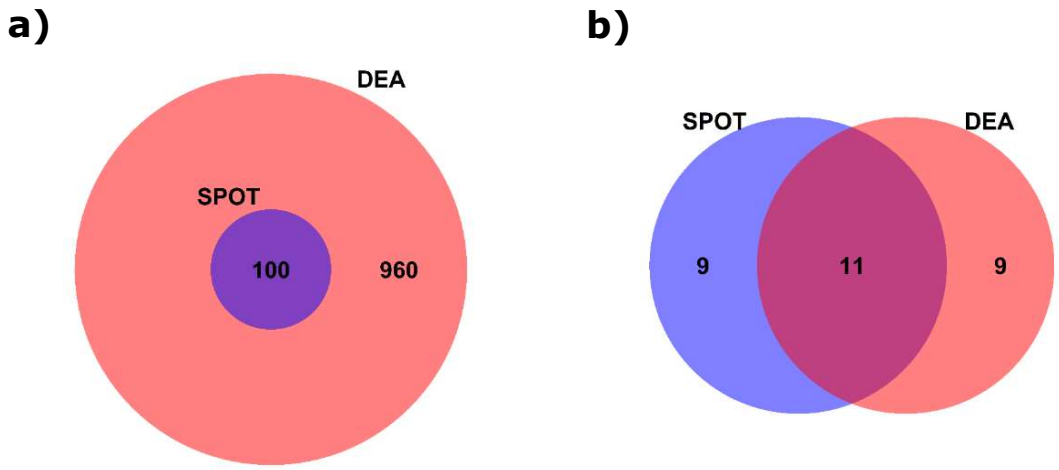

Figure S7: The majority of top SPOT sex specific genes overlap with DEA results

(a) All genes ranked in the *spot* Top 100 overlap with genes detected as significantly upregulated in the sexual stages (1060). **(b)** Comparison of top 20 genes derived from *spot* or the Wilcoxon Rank Sum test (DEA, ranked by p-value) reveal an overlap of 11 genes.

While the genes in both rankings have very low values in the unselected stages, there are substantially higher expression values in selected stages of the genes ranked by *spot* compared to those ranked by Wilcoxon (Figure S8).

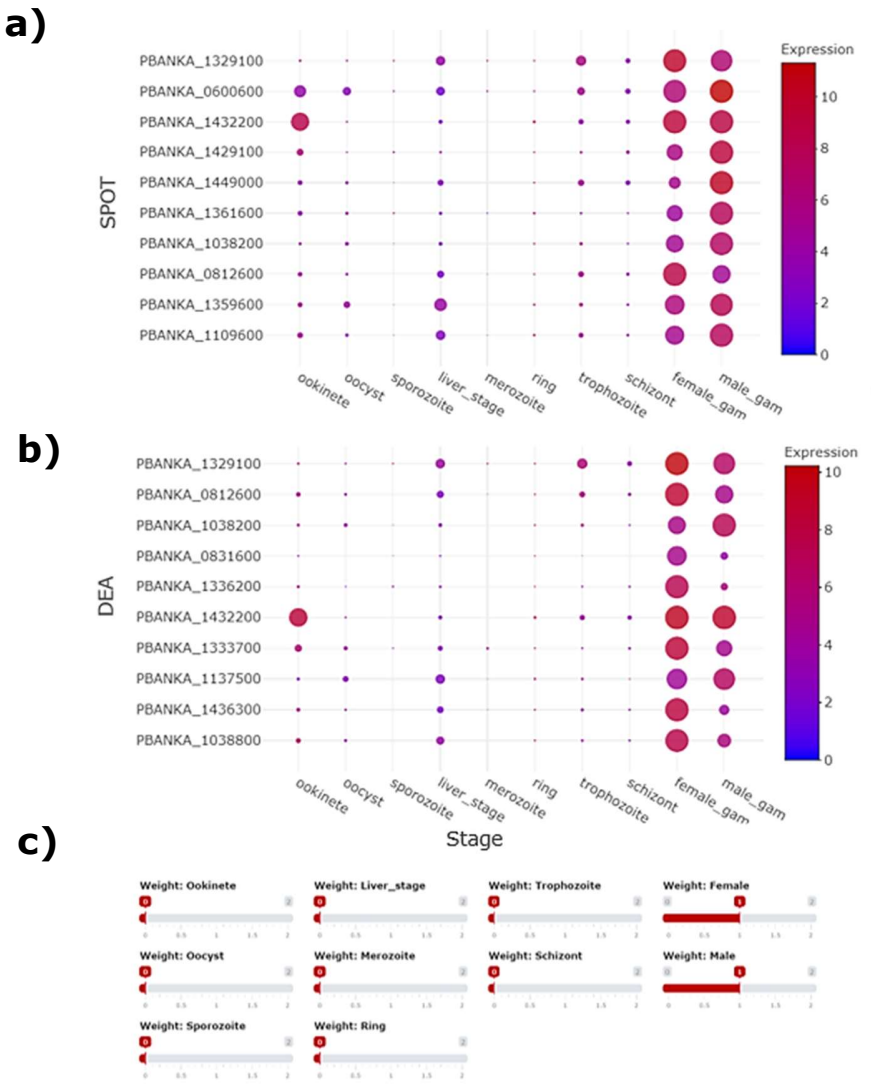

Figure S8: *spot* ranked genes reveal high expression in sexual stages

Genes ranking high in the *spot* algorithm **(a)** have substantially higher expression values in sexual stages than Wilcoxon ranked genes shown in **(b).** Similar to the liver stage results (Figure S6) genes ranking high in the Wilcoxon test have slightly lower expression values in the unselected columns than genes ranked by *spot*. **(c)** The slider constellation of SPOT to obtain the results shown in **(a)** and **(b)**.

5 Example prediction C: Genes expressed in asexual and sexual blood stages of *Plasmodium*

5.1 Purpose

Asexual and sexual blood stages share the red cell as a host, yet their biology differs dramatically as has also been shown by shifting expression profiles. Here we present a prediction of genes only active in asexual or sexual blood stages but not in other life cycle stages. This should counter select genes involved in general replication mechanisms.

5.2 Results

The results of the third example prediction shown in Table S4 display a multitude of membrane proteins or proteins associated with membrane trafficking and virulence. While the membrane associated histidine-rich protein 1a (MAHRP1a) and the skeleton-binding protein 1 (SBP1) are known to play a role in the transport of parasite proteins to the surface (Blisnick, et al., 2000; De Niz, et al., 2016; Maier, et al., 2007; Spycher, et al., 2003), the function of the ETRAMP protein family is more diverse and needs further evaluation (MacKellar, et al., 2011; Spielmann, et al., 2003). This also applies to the fam-b protein and the p1/s1 nuclease, about which little is known, as well as to the two unknown proteins that are candidates for further analysis.

Table S4: Results from *spot* ranking of blood and sexual stage specific genes

P. Berghei GeneID	Gene product description	*spot* score
PBANKA_1145800	membrane associated histidine-rich protein 1a	5,61
PBANKA_1000600	erythrocyte membrane antigen 1	4,96
PBANKA_0300600	Plasmodium exported protein, unknown function	4,78
PBANKA_0524800	early transcribed membrane protein	4,64
PBANKA_1101300	skeleton-binding protein 1	4,52
PBANKA_0517000	early transcribed membrane protein	4,13
PBANKA_0316300	fam-b protein	3,62
PBANKA_1200600	Plasmodium exported protein, unknown function	3,51
PBANKA_0524200	early transcribed membrane protein	3,47
PBANKA_1030600	p1/s1 nuclease, putative	3,47

5.3 Comparison to DEA

In contrast to the previous predictions, a substantially lower number of genes was significantly upregulated according to the Wilcoxon test (Figure S9 a). This might reflect the fact that many genes needed for proliferation will also be expressed in oocysts and liver stages. Nevertheless, the large majority of genes derived from *spot* ranking still overlapped with these genes. The overlap between the top 20 genes from Wilcoxon and *spot* ranking is 50 %.

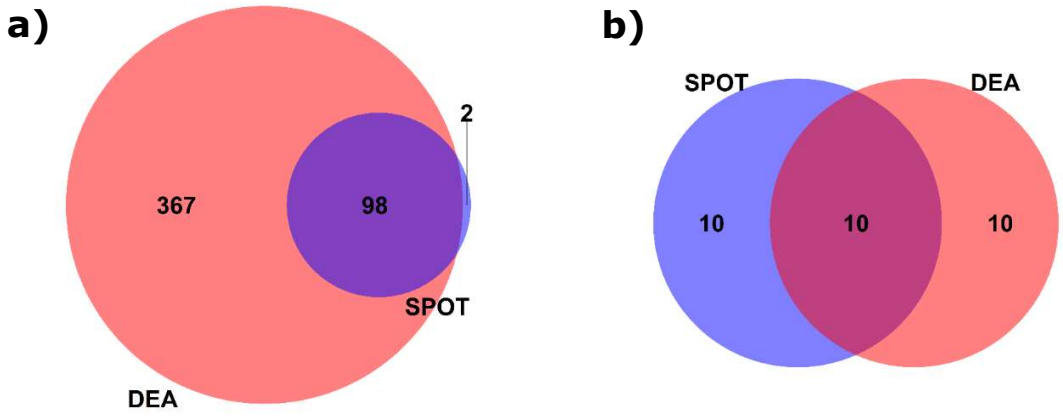

Figure S9: In more complex searches results from *spot* ranking are still differentially expressed

(a) 98 of 100 genes ranked in the *spot* top 100 overlap with genes detected as significantly upregulated in the asexual and sexual blood stages (465). 465 is by far the lowest number of significantly expressed genes in the example predictions. **(b)** 50 % of top 20 genes derived from *spot* ranking or the Wilcoxon Rank Sum test (DEA, ranked by p-value) overlap.

However, the genes scoring best in the Wilcoxon test ranking show very high expression in unselected stages compared to all other predictions (Figure S4; Figure S10 b). While the expression values from *spot* ranking derived genes stay low compared to other predictions, they show relatively low expression in the selected stages (Figure S4; Figure S10 a)).

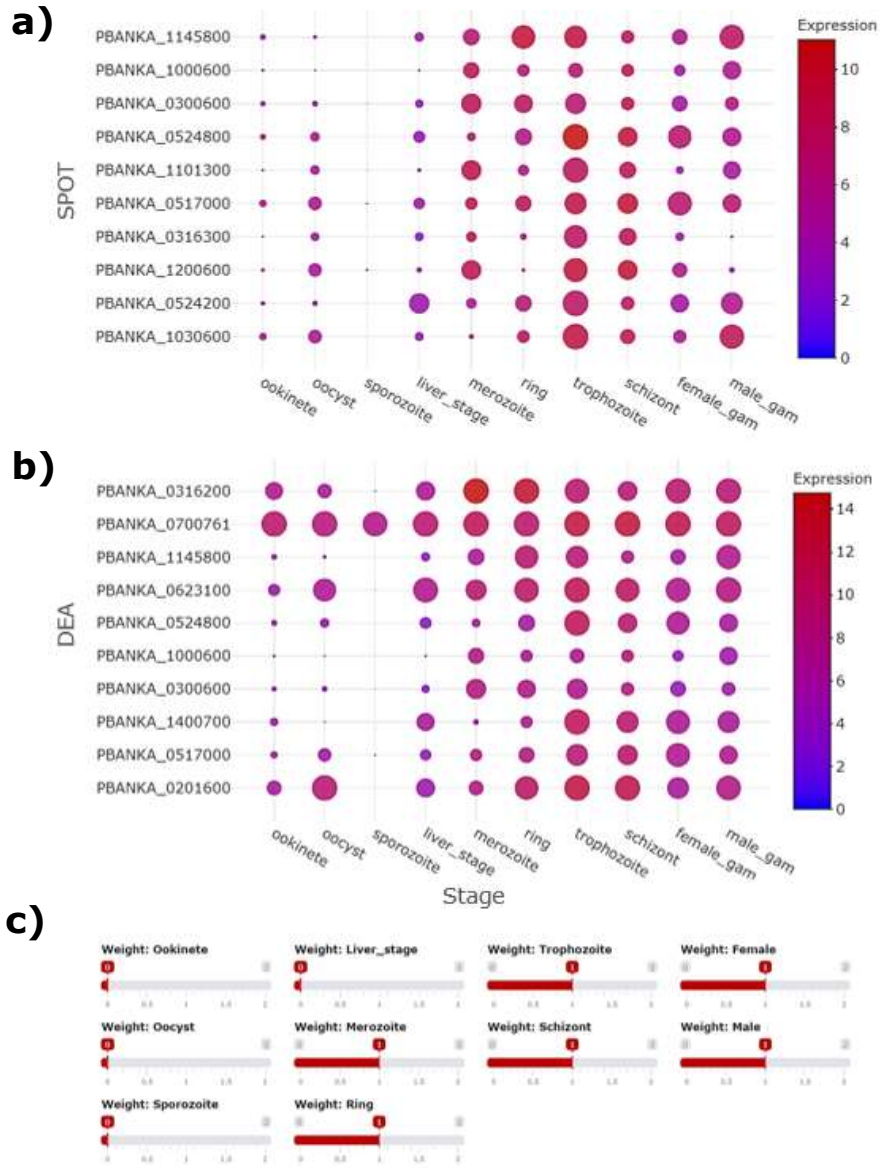

Figure S10: *spot* ranking for transmission targets shows low expression in unselected stages

(a) Genes ranking in the top 10 of *spot* ranking have substantially lower gene expression values in unselected columns compared to genes ranking in the top 10 in the Wilcoxon test **(b)**. However, Wilcoxon test ranked genes shown in **(b)** have substantially higher expression values in selected stages. Additionally, genes derived by the Wilcoxon test show a higher variance in both categories. **(c)** The slider constellation of SPOT to obtain the results shown in **(a)** and **(b)**.

6 Requirements

Table S5: Packages used for web-tool

Package	Version
DESeq2	1.30.1
DT	0.18
edgeR	3.32.1
EnvStats	2.4.0
limma	3.46.0
MAST	1.16.0
matrixStats	0.58.0
plotly	4.9.3
plyr	1.8.6
reshape2	1.4.4
rlist	0.4.6.1
scales	1.1.1
Seurat	4.0.1
shiny	1.6.0
shinybusy	0.2.2
shinycssloaders	1.0.0
shinyEffects	0.2.0
shinyWidgets	0.6.0
sortable	0.4.4
stringr	1.4.0
tidyr	1.1.3

7 Acknowledgements

We thank Omar Harb, Daniel Beiting and the reviewers for critical reading and constructive suggestions on the manuscript as well as Benedikt Broers and Felix Borchers for discussions.

8 References of supplement

Acquah, F.K., et al. (2019). Transmission-Blocking Vaccines: Old Friends and New Prospects. Infect. Immun. 87, e00775-00718.

Blisnick, T., et al. (2000). Pfsbp1, a Maurer's cleft Plasmodium falciparum protein, is associated with the erythrocyte skeleton. Mol Biochem Parasitol 111, 107-121.

Bushell, E.S.C., et al. (2009). Paternal Effect of the Nuclear Formin-like Protein MISFIT on Plasmodium Development in the Mosquito Vector. PLoS Pathog 5, e1000539.

Caldelari, R., et al. (2019). Transcriptome analysis of Plasmodium berghei during exo-erythrocytic development. Malaria J 18, 330-320.

Cardoso-Moreira, M., et al. (2019). Gene expression across mammalian organ development. Nature 571, 505-509.

Carter, R. and Chen, D.H. (1976). Malaria transmission blocked by immunisation with gametes of the malaria parasite. Nature 263, 57-60.

De Niz, M., et al. (2016). The machinery underlying malaria parasite virulence is conserved between rodent and human malaria parasites. Nat Commun 7, 11659.

Deligianni, E., et al. (2018). Sequence and functional divergence of gametocyte-specific parasitophorous vacuole membrane proteins in Plasmodium parasites. Mol Biochem Parasitol 220, 15-18.

Duffy, P.E. and Patrick Gorres, J. (2020). Malaria vaccines since 2000: progress, priorities, products. NPJ Vaccines 5, 48.

Flannery, E.L., et al. (2018). Assessing drug efficacy against Plasmodium falciparum liver stages in vivo. JCI insight 3, e92587.

Ge, S.X. et al. (2018). iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. BMC Bioinformatics 19, 534.

Gwadz, R.W. (1976). Successful immunization against the sexual stages of Plasmodium gallinaceum. Science 193, 1150-1151.

Howick, V.M., et al. (2019). The Malaria Cell Atlas: Single parasite transcriptomes across the complete Plasmodium life cycle. Science 365, 56.

Ishino, T., et al. (2009). LISP1 is important for the egress of Plasmodium berghei parasites from liver cells. Cell Microbiol 11, 1329-1339.

Jiang, Y., et al. (2020). An intracellular membrane protein GEP1 regulates xanthurenic acid induced gametogenesis of malaria parasites. Nat Commun 11, 1764.

Kaiser, K., et al. (2004). Differential transcriptome profiling identifies Plasmodium genes encoding pre-erythrocytic stage-specific proteins. Mol Microbiol 51, 1221-1232.

Kappe, S.H.I., et al. (2010). That Was Then But This Is Now: Malaria Research in the Time of an Eradication Agenda. Science 328, 862.

Kehrer, J., et al. (2016). Proteomic Analysis of the Plasmodium berghei Gametocyte Egressome and Vesicular bioID of Osmiophilic Body Proteins Identifies Merozoite TRAP-like Protein (MTRAP) as an Essential Factor for Parasite Transmission. Mol Cell Proteom 15, 2852.

Kumar, H., et al. (2016). Protective efficacy and safety of liver stage attenuated malaria parasites. Sci Rep 6, 26824.

Lee, S.-M., et al. (2020). A C-terminal Pfs48/45 malaria transmission-blocking vaccine candidate produced in the baculovirus expression system. Sci Rep 10, 395.

Love M.I., et al. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550.

MacKellar, D.C., et al. (2011). A systematic analysis of the early transcribed membrane protein family throughout the life cycle of Plasmodium yoelii. Cell Microbiol 13, 1755-1767.

Maier, A.G., et al. (2007). Skeleton-binding protein 1 functions at the parasitophorous vacuole membrane to traffic PfEMP1 to the Plasmodium falciparum-infected erythrocyte surface. Blood 109, 1289-1297.

Matuschewski, K., et al. (2002). Infectivity-associated changes in the transcriptional repertoire of the malaria parasite sporozoite stage. J Biol Chem 277, 41948-41953.

Mueller, A.-K., et al. (2005). Genetically modified Plasmodium parasites as a protective experimental malaria vaccine. Nature 433, 164-167.

Orito, Y., et al. (2003). Liver-specific protein 2: a Plasmodium protein exported to the hepatocyte cytoplasm and required for merozoite formation. Mol Microbiol 87, 66-79.

Reyes, A.L.P., et al. (2019). GENAVi: a shiny web application for gene expression normalization, analysis and visualization. BMC Genomics 20, 745.

Ponzi, M., et al. (2009). Egress of Plasmodium berghei gametes from their host erythrocyte is mediated by the MDV-1/PEG3 protein. Cell Microbiol 11, 1272-1288.

RTS,S Clinical Trials Partnership (2015). Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomised, controlled trial. Lancet 386, 31-45.

Soneson, C. and Robinson, M.D. (2018). Bias, robustness and scalability in single-cell differential expression analysis. Nat Methods 15, 255-261.

Spielmann, T., (2003). etramps, a new Plasmodium falciparum gene family coding for developmentally regulated and highly charged membrane proteins located at the parasite-host cell interface. Mol Biol Cell 14, 1529-1544.

Spycher, C., et al. (2003). MAHRP-1, a novel Plasmodium falciparum histidine-rich protein, binds ferriprotoporphyrin IX and localizes to the Maurer's clefts. J Biol Chem 278, 35373-35383.

Stanway, R.R., et al. (2019). Genome-Scale Identification of Essential Metabolic Processes for Targeting the Plasmodium Liver Stage. Cell 179, 1112-1128 e1126.

The malERA Consultative Group on Drugs (2011). A Research Agenda for Malaria Eradication: Drugs. PLoS Med 8, e1000402.

van Dijk, M.R., et al. (2001). A central role for P48/45 in malaria parasite male gamete fertility. Cell 104, 153-164.

Vaughan, A.M., et al. (2009). Type II fatty acid synthesis is essential only for malaria parasite late liver stage development. Cell Microbiol 11, 506-520.

Weißbach, T., et al. (2017). Transcript and protein expression analysis of proteases in the blood stages of Plasmodium falciparum. Exp Parasitol 180, 33-44.

Wyler, E., et al. (2021). Transcriptomic profiling of SARS-CoV-2 infected human cell lines identifies HSP90 as target for COVID-19 therapy. iScience 24, 102151.